

ON THE BOUNDEDNESS OF BI-PARAMETER LITTLEWOOD-PALEY g_λ^* -FUNCTION

MINGMING CAO AND QINGYING XUE

ABSTRACT. Let $m, n \geq 1$ and $g_{\lambda_1, \lambda_2}^*$ be the bi-parameter Littlewood-Paley g_λ^* -function defined by

$$g_{\lambda_1, \lambda_2}^*(f)(x) = \left(\iint_{\mathbb{R}_+^{m+1}} \left(\frac{t_2}{t_2 + |x_2 - y_2|} \right)^{m\lambda_2} \iint_{\mathbb{R}_+^{n+1}} \left(\frac{t_1}{t_1 + |x_1 - y_1|} \right)^{n\lambda_1} \right. \\ \left. \times |\theta_{t_1, t_2} f(y_1, y_2)|^2 \frac{dy_1 dt_1}{t_1^{n+1}} \frac{dy_2 dt_2}{t_2^{m+1}} \right)^{1/2}, \quad \lambda_1 > 1, \quad \lambda_2 > 1$$

where $\theta_{t_1, t_2} f$ is a non-convolution kernel defined on \mathbb{R}^{m+n} . In this paper, we showed that the bi-parameter Littlewood-Paley function $g_{\lambda_1, \lambda_2}^*$ was bounded from $L^2(\mathbb{R}^{n+m})$ to $L^2(\mathbb{R}^{n+m})$. This was done by means of probabilistic methods and by using a new averaging identity over good double Whitney regions.

1. INTRODUCTION

1.1. Background and motivation. It is well known that g_λ^* -function originated in the work of Littlewood and Paley [9] in the 1930's. In 1961, Stein [16] introduced and studied the following higher dimensional ($n \geq 2$) Littlewood-Paley g_λ^* -function:

$$g_\lambda^*(f)(x) = \left(\iint_{\mathbb{R}_+^{n+1}} \left(\frac{t}{t + |x - y|} \right)^{n\lambda} |\nabla P_t f(y, t)|^2 \frac{dy dt}{t^{n-1}} \right)^{1/2}, \quad \lambda > 1$$

where $P_t f(y, t) = p_t * f(x)$, $p_t(y) = t^{-n} p(y/t)$ denotes the Poisson kernel and $\nabla = (\frac{\partial}{\partial y_1}, \dots, \frac{\partial}{\partial y_n}, \frac{\partial}{\partial t})$. It plays important roles in harmonic analysis and other fields. With much greater difficulty, Stein [15] showed that $\|g_\lambda^*(f)\|_{L^p(\mathbb{R}^n)}$ and $\|f\|_{L^p(\mathbb{R}^n)}$ are equivalent norms for any $1 < p < \infty$. Moreover, in [15], Stein also proved that if $\lambda > 2$, then g_λ^* is of weak type $(1, 1)$, and is of strong type (p, p) for $1 < p < \infty$. In the same paper, Stein pointed out that weak $(1, 1)$ boundedness doesn't hold for $1 < \lambda \leq 2$. In 1970, as a replacement of weak $(1, 1)$ bounds for $1 < \lambda < 2$, Fefferman [3] considered the end-point weak (p, p) estimates of g_λ^* -function when $p > 1$ and $\lambda = 2/p$.

Recently, Cao, Xue and Li [1] gave a characterization of two weight norm inequalities for the classical g_λ^* -function. The first step of the proof is to reduce the case to good

Date: December 1, 2015.

Key words and phrases. Bi-parameter Littlewood-Paley g_λ^* -function; Probabilistic methods; Haar functions.

The second author was supported partly by NSFC (No. 11471041), the Fundamental Research Funds for the Central Universities (NO. 2012CXQT09 and NO. 2014KJJCA10) and NCET-13-0065.

Corresponding author: Qingying Xue Email: qyxue@bnu.edu.cn.

Whitney regions. In addition, the random dyadic grids and martingale differences decomposition are used. The core of the proof is the construction of stopping cubes, which is a modern and effective technique to deal with two weight problems. The stopping cubes were first introduced to handle two weight boundedness of Hilbert transform [6], [7]. Then, some related consequences and applications were given, as demonstrated in [1], [8] and [10]. Still more recently, Cao and Xue [2] established a local Tb theorem for the non-homogeneous Littlewood-Paley g_λ^* -function with non-convolution type kernels and upper power bound measure μ . It was the first time to investigate g_λ^* -function in the simultaneous presence of three attributes : local, non-homogeneous and L^p -testing condition.

When it comes to the multi-parameter harmonic analysis, there is a very large existing theory. In 2012, a dyadic representation theorem for bi-parameter singular integrals was presented by Martikainen [11] and a new version of $T1$ theorem on the product space was also established. In 2014, Hytönen and Martikainen [5] proved a non-homogeneous version of $T1$ theorem for certain bi-parameter singular integral operators. Moreover, they discussed the related non-homogeneous Journé's lemma and product BMO theory with more general type of measures. Still in 2014, a class of bi-parameter kernels and related vertical square functions in the upper half-space were first introduced by Martikainen [12]. Using dyadic probabilistic techniques, the author gave a criterion for the $L^2(\mathbb{R}^{n+m})$ boundedness of these square functions. It is worth pointing out that the kernels are assumed to satisfy some estimates, including a natural size condition, a Hölder estimate and two symmetric mixed Hölder and size estimates, the mixed Carleson and size conditions, the mixed Carleson and Hölder estimates and a bi-parameter Carleson condition. Moreover, it should be noted that the bi-parameter Carleson condition is necessary for the square function to be bounded on $L^2(\mathbb{R}^{n+m})$.

Motivated by the above works, in this paper, we keep on studying the Littlewood Paley g_λ^* -function but in the bi-parameter setting. First, we introduce the definition of the bi-parameter Littlewood Paley g_λ^* -function.

Definition 1.1. Let $\lambda_1, \lambda_2 > 1$, for any $x = (x_1, x_2) \in \mathbb{R}^{n+m}$, the bi-parameter Littlewood-Paley g_λ^* -function is defined by

$$g_{\lambda_1, \lambda_2}^*(f)(x) := \left(\iint_{\mathbb{R}_+^{m+1}} \left(\frac{t_2}{t_2 + |x_2 - y_2|} \right)^{m\lambda_2} \iint_{\mathbb{R}_+^{n+1}} \left(\frac{t_1}{t_1 + |x_1 - y_1|} \right)^{n\lambda_1} \right. \\ \left. \times |\theta_{t_1, t_2} f(y_1, y_2)|^2 \frac{dy_1 dt_1}{t_1^{n+1}} \frac{dy_2 dt_2}{t_2^{m+1}} \right)^{1/2},$$

where $\theta_{t_1, t_2} f(y_1, y_2) = \iint_{\mathbb{R}^{n+m}} K_{t_1, t_2}(y_1, y_2, z_1, z_2) f(z_1, z_2) dz_1 dz_2$.

Under certain structural assumptions, we will prove the $L^2(\mathbb{R}^{n+m})$ boundedness of $g_{\lambda_1, \lambda_2}^*$, in other words, the following inequality,

$$\|g_{\lambda_1, \lambda_2}^*(f)\|_{L^2(\mathbb{R}^{n+m})} \lesssim \|f\|_{L^2(\mathbb{R}^{n+m})}.$$

Compared to the bi-parameter vertical square function, the bi-parameter Littlewood Paley g_λ^* -function is significantly much more difficult to be dealt with. Actually, in bi-parameter case, additional integrals make most of the corresponding estimates more complicated. We could not use the assumptions in [12] directly, since addition terms

appears in the Definition 1.1. In fact, we will use much more weaker conditions than the conditions used in [12] (see assumptions in the following subsection). Unlike the one-parameter case and two-weight case [1], the proof of bi-parameter g_λ^* -function does not involve the stopping cubes and martingale differences decomposition. In fact, the decomposition associated with Haar function in \mathbb{R}^n provides a foundation for our analysis. And modern techniques, including probabilistic methods and dyadic analysis, will be used efficiently again. They were first used by Martikainen [11] in the study of the bi-parameter Calderón-Zygmund integrals and later appeared in [12]. For more applications, one can refer [5], [14]. However, our result is based on a simple new averaging identity over good double Whitney regions.

1.2. Assumptions and Main result. To state our main results, we need to give some appropriate assumptions. From now on, we always assume that $\alpha, \beta > 0$. We use, for minor convenience, ℓ^∞ metrics on \mathbb{R}^n and \mathbb{R}^m .

Assumption 1 (Standard estimates). The kernel $K_{t_1, t_2} : \mathbb{R}^{n+m} \times \mathbb{R}^{n+m} \rightarrow \mathbb{C}$ is assumed to satisfy the following estimates:

(1) Size condition :

$$|K_{t_1, t_2}(x, y)| \lesssim \frac{t_1^\alpha}{(t_1 + |x_1 - y_1|)^{n+\alpha}} \frac{t_2^\beta}{(t_2 + |x_2 - y_2|)^{m+\beta}}.$$

(2) Hölder condition :

$$\begin{aligned} & |K_{t_1, t_2}(x, y) - K_{t_1, t_2}(x, (y_1, y'_2)) - K_{t_1, t_2}(x, (y'_1, y_2)) + K_{t_1, t_2}(x, y')| \\ & \lesssim \frac{|y_1 - y'_1|^\alpha}{(t_1 + |x_1 - y_1|)^{n+\alpha}} \frac{|y_2 - y'_2|^\beta}{(t_2 + |x_2 - y_2|)^{m+\beta}}, \end{aligned}$$

whenever $|y_1 - y'_1| < t_1/2$ and $|y_2 - y'_2| < t_2/2$.

(3) Mixed Hölder and size conditions :

$$|K_{t_1, t_2}(x, y) - K_{t_1, t_2}(x, (y_1, y'_2))| \lesssim \frac{t_1^\alpha}{(t_1 + |x_1 - y_1|)^{n+\alpha}} \frac{|y_2 - y'_2|^\beta}{(t_2 + |x_2 - y_2|)^{m+\beta}},$$

whenever $|y_2 - y'_2| < t_2/2$ and

$$|K_{t_1, t_2}(x, y) - K_{t_1, t_2}(x, (y'_1, y_2))| \lesssim \frac{|y_1 - y'_1|^\alpha}{(t_1 + |x_1 - y_1|)^{n+\alpha}} \frac{t_2^\beta}{(t_2 + |x_2 - y_2|)^{m+\beta}},$$

whenever $|y_1 - y'_1| < t_1/2$.

Assumption 2 (Carleson condition \times Standard estimates). If $I \subset \mathbb{R}^n$ is a cube with side length $\ell(I)$, we define the associated Carleson box by $\widehat{I} = I \times (0, \ell(I))$. We assume the following conditions : For every cube $I \subset \mathbb{R}^n$ and $J \subset \mathbb{R}^m$, there holds that

(1) Combinations of Carleson and size conditions :

$$\begin{aligned} & \left(\iint_{\widehat{I}} \int_{\mathbb{R}^n} \left| \int_I K_{t_1, t_2}(x - y, z_1, z_2) dz_1 \right|^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \frac{dy_1 dx_1 dt_1}{t_1^{n+1}} \right)^{\frac{1}{2}} \\ & \lesssim |I|^{\frac{1}{2}} \frac{t_2^\beta}{(t_2 + |x_2 - y_2 - z_2|)^{m+\beta}} \end{aligned}$$

and

$$\begin{aligned} & \left(\iint_{\widehat{J}} \int_{\mathbb{R}^m} \left| \int_J K_{t_1, t_2}(x - y, z_1, z_2) dz_2 \right|^2 \left(\frac{t_2}{t_2 + |y_2|} \right)^{n\lambda_1} \frac{dy_2 dx_2 dt_2}{t_2^{m+1}} \right)^{\frac{1}{2}} \\ & \lesssim |J|^{\frac{1}{2}} \frac{t_1^\alpha}{(t_1 + |x_1 - y_1 - z_1|)^{n+\alpha}}. \end{aligned}$$

(2) Combinations of Carleson and Hölder conditions :

$$\begin{aligned} & \left(\iint_{\widehat{I}} \int_{\mathbb{R}^n} \left| \int_I [K_{t_1, t_2}(x - y, z_1, z_2) - K_{t_1, t_2}(x - y, z_1, z'_2)] dz_1 \right|^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \frac{dy_1 dx_1 dt_1}{t_1^{n+1}} \right)^{\frac{1}{2}} \\ & \lesssim |I|^{\frac{1}{2}} \frac{|z_2 - z'_2|^\beta}{(t_2 + |x_2 - y_2 - z_2|)^{m+\beta}}, \end{aligned}$$

whenever $|z_2 - z'_2| < t_2/2$. And

$$\begin{aligned} & \left(\iint_{\widehat{J}} \int_{\mathbb{R}^m} \left| \int_J [K_{t_1, t_2}(x - y, z_1, z_2) - K_{t_1, t_2}(x - y, z'_1, z_2)] dz_2 \right|^2 \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_2 dx_2 dt_2}{t_2^{m+1}} \right)^{\frac{1}{2}} \\ & \lesssim |J|^{\frac{1}{2}} \frac{|z_1 - z'_1|^\alpha}{(t_1 + |x_1 - y_1 - z_1|)^{n+\alpha}}, \end{aligned}$$

whenever $|z_1 - z'_1| < t_1/2$.

Assumption 3 (Bi-parameter Carleson condition). Let $\mathcal{D} = \mathcal{D}_n \times \mathcal{D}_m$, where \mathcal{D}_n is a dyadic grid in \mathbb{R}^n and \mathcal{D}_m is a dyadic grid in \mathbb{R}^m . For $I \in \mathcal{D}_n$, let $W_I = I \times (\ell(I)/2, \ell(I))$ be the associated Whitney region. Denote $n_1 = n, n_2 = m$ and

$$C_{IJ}^{\mathcal{D}} = \iint_{W_J} \iint_{W_I} \iint_{\mathbb{R}^{n+m}} |\theta_{t_1, t_2} \mathbf{1}(y_1, y_2)|^2 \left[\prod_{i=1}^2 \left(\frac{t_i}{t_i + |x_i - y_i|} \right)^{n_i \lambda_i} \right] \frac{dy_1 dy_2}{t_1^n t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2}.$$

We assume the following *bi-parameter Carleson condition*: For every $\mathcal{D} = \mathcal{D}_n \times \mathcal{D}_m$ there holds that

$$(1.1) \quad \sum_{\substack{I \times J \in \mathcal{D} \\ I \times J \subset \Omega}} C_{IJ}^{\mathcal{D}} \lesssim |\Omega|$$

for all sets $\Omega \subset \mathbb{R}^{n+m}$ such that $|\Omega| < \infty$ and such that for every $x \in \Omega$ there exists $I \times J \in \mathcal{D}$ so that $x \in I \times J \subset \Omega$.

We now formulate the main result of this paper as follows.

Theorem 1.1. *Let $\lambda_1, \lambda_2 > 2$, $0 < \alpha \leq n(\lambda_1 - 2)/2$ and $0 < \beta \leq m(\lambda_2 - 2)/2$. Assume that the kernel K_{t_1, t_2} satisfies the Assumptions 1-3. Then there holds that*

$$\|g_{\lambda_1, \lambda_2}^*(f)\|_{L^2(\mathbb{R}^{n+m})} \lesssim \|f\|_{L^2(\mathbb{R}^{n+m})},$$

where the implied constant depends only on the assumptions.

Remark 1.2. In section 6, we shall show that the bi-parameter Carleson condition is necessary for $g_{\lambda_1, \lambda_2}^*$ -function to be bounded on $L^2(\mathbb{R}^{n+m})$. Moreover, Assumption 2 and Assumption 3 are much weaker than the similar conditions used in [12], since here two terms (both less than one) were added and more integrals related to y_1 or y_2 were used in our assumptions.

2. THE PROBABILISTIC REDUCTION

In this section, our goal is to simplify the proof of the main result. First, we recall the definitions of random dyadic grids, good/bad cubes, Haar function on \mathbb{R}^n which can be found in [4], [11] and [13].

2.1. Random Dyadic Grids. Let $\beta_n = \{\beta_n^j\}_{j \in \mathbb{Z}}$, where $\beta_n^j \in \{0, 1\}^n$. Let \mathcal{D}_n^0 be the standard dyadic grids on \mathbb{R}^n . We define the new dyadic grids in \mathbb{R}^n by

$$\mathcal{D}_n = \left\{ I + \beta_n; I \in \mathcal{D}_n^0 \right\} := \left\{ I + \sum_{j: 2^{-j} < \ell(I)} 2^{-j} \beta_n^j; I \in \mathcal{D}_n^0 \right\}.$$

Similarly, we can define the dyadic grids \mathcal{D}_m in \mathbb{R}^m . There is a natural product probability structure on $(\{0, 1\}^n)^\mathbb{Z}$ and $(\{0, 1\}^m)^\mathbb{Z}$. Therefore, we have independent random dyadic grids \mathcal{D}_n and \mathcal{D}_m , respectively.

2.2. Good and Bad Cubes. A cube $I \in \mathcal{D}_n$ is said to be *bad* if there exists a $J \in \mathcal{D}_n$ with $\ell(J) \geq 2^r \ell(I)$ such that $\text{dist}(I, \partial J) \leq \ell(I)^{\gamma_n} \ell(J)^{1-\gamma_n}$. Otherwise, I is called *good*. Here $r \in \mathbb{Z}_+$ and $\gamma_n \in (0, \frac{1}{2})$ are given parameters. Roughly speaking, a dyadic cube I will be bad if it is relatively close to the boundary of a much bigger dyadic cube. Denote $\pi_{\text{good}}^n = \mathbb{P}_{\beta_n}(I + \beta_n \text{ is good}) = \mathbb{E}_{\beta_n}(\mathbf{1}_{\text{good}}(I + \beta_n))$. Then π_{good}^n is independent of $I \in \mathcal{D}_n^0$, and the parameter r is a fixed constant so that $\pi_{\text{good}}^n, \pi_{\text{good}}^m > 0$.

Throughout this article, we take $\gamma_n = \frac{\alpha}{2(n+\alpha)}$, where $\alpha > 0$ appears in the kernel estimates. It is important to observe that the position and goodness of a cube $I \in \mathcal{D}_n^0$ are independent. Indeed, according to the definition, the spatial position of

$$I + \beta_n = I + \sum_{j: 2^{-j} < \ell(I)} 2^{-j} \beta_n^j$$

depends only on β_n^j for $2^{-j} < \ell(I)$. On the other hand, the relative position of $I + \beta_n$ with respect to a bigger cube

$$J + \beta_n = I + \sum_{j: 2^{-j} < \ell(I)} 2^{-j} \beta_n^j + \sum_{j: \ell(I) \leq 2^{-j} < \ell(J)} 2^{-j} \beta_n^j$$

depends only on β_n^j for $\ell(I) \leq 2^{-j} < \ell(J)$. Thus, the position and goodness of $I + \beta_n$ are independent.

2.3. Haar functions. In order to decompose a function $f \in L^2$, we introduce the definition of the Haar functions on \mathbb{R}^n . Let h_I be an L^2 normalized Haar function related to $I \in \mathcal{D}_n$, where \mathcal{D}_n is a dyadic grid on \mathbb{R}^n . With this we mean that $h_I, I = I_1 \times \cdots \times I_n$, is one of the 2^n functions h_I^η , $\eta = \eta_1, \dots, \eta_n \in \{0, 1\}^n$, defined by

$$h_I^\eta = h_{I_1}^{\eta_1} \otimes \cdots \otimes h_{I_n}^{\eta_n},$$

where $h_{I_i}^0 = |I_i|^{-1/2} \mathbf{1}_{I_i}$ and $h_{I_i}^1 = |I_i|^{-1/2} (\mathbf{1}_{I_{i,l}} - \mathbf{1}_{I_{i,r}})$ for every $i = 1, \dots, n$. Here $I_{i,l}$ and $I_{i,r}$ are the left and right halves of the interval I_i respectively. If $\eta \neq 0$, the Haar function is cancellative : $\int_{\mathbb{R}^n} h_I = 0$. All the cancellative Haar functions form an orthonormal basis of $L^2(\mathbb{R}^n)$. If $a \in L^2(\mathbb{R}^n)$, we may thus write

$$a = \sum_{I \in \mathcal{D}_n} \sum_{\eta \in \{0,1\}^n \setminus \{0\}} \langle a, h_I^\eta \rangle h_I^\eta.$$

However, we suppress the finite η summation and just write $a = \sum_I \langle a, h_I \rangle h_I$. Using the corresponding product basis, we may expand a function f defined in \mathbb{R}^{n+m} in the following way:

$$f = \sum_{I,J} f_{IJ} h_{I \times J} := \sum_{I,J} \langle f, h_I \otimes h_J \rangle h_I \otimes h_J.$$

2.4. Averaging over Good Whitney Regions. Let $f \in L^2(\mathbb{R}^{n+m})$. Let $I_1, I_2 \in \mathcal{D}_n$ and $J_1, J_2 \in \mathcal{D}_m$. Note that the position and goodness of $I + \beta_n$ are independent. Therefore, one can write,

$$\begin{aligned} & \|g_{\lambda_1, \lambda_2}^*(f)\|_{L^2(\mathbb{R}^{n+m})}^2 \\ &= \iint_{\mathbb{R}_+^{m+1}} \iint_{\mathbb{R}_+^{n+1}} \iint_{\mathbb{R}^{n+m}} |\theta_{t_1, t_2} f(x-y)|^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \\ & \quad \times \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2} \\ &= \frac{1}{\pi_{good}^n} \frac{1}{\pi_{good}^m} \mathbb{E}_{\beta_n, \beta_m} \sum_{I_2, J_2: good} \iint_{W_{J_2}} \iint_{W_{I_2}} \iint_{\mathbb{R}^{n+m}} |\theta_{t_1, t_2} f(x-y)|^2 \\ & \quad \times \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2} \\ &= \frac{1}{\pi_{good}^n} \frac{1}{\pi_{good}^m} \mathbb{E}_{\beta_n, \beta_m} \sum_{I_2, J_2: good} \iint_{W_{J_2}} \iint_{W_{I_2}} \iint_{\mathbb{R}^{n+m}} \left| \sum_{I_1, J_1} f_{I_1 J_1} \theta_{t_1, t_2} h_{I_1 \times J_1}(x-y) \right|^2 \\ & \quad \times \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2}. \end{aligned}$$

Consequently, we are reduced to bound the sum

$$\begin{aligned} \mathcal{G} &:= \sum_{I_2, J_2: good} \iint_{W_{J_2}} \iint_{W_{I_2}} \iint_{\mathbb{R}^{n+m}} \left| \sum_{I_1, J_1} f_{I_1 J_1} \theta_{t_1, t_2} h_{I_1 \times J_1}(x-y) \right|^2 \\ & \quad \times \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2}. \end{aligned}$$

Furthermore, we can carry out the decomposition

$$\mathcal{G} \lesssim \mathcal{G}_{<, <} + \mathcal{G}_{<, \geq} + \mathcal{G}_{\geq, <} + \mathcal{G}_{\geq, \geq},$$

where

$$\begin{aligned} \mathcal{G}_{<, <} &:= \sum_{I_2, J_2: good} \iint_{W_{J_2}} \iint_{W_{I_2}} \iint_{\mathbb{R}^{n+m}} \left| \sum_{\substack{I_1, J_1 \\ \ell(I_1) < \ell(I_2) \\ \ell(J_1) < \ell(J_2)}} f_{I_1 J_1} \theta_{t_1, t_2} h_{I_1 \times J_1}(x-y) \right|^2 \\ & \quad \times \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2}, \end{aligned}$$

and the others are completely similar.

Sequentially, it is enough to focus on estimating the four pieces: $\mathcal{G}_{<, <}$, $\mathcal{G}_{<, \geq}$, $\mathcal{G}_{\geq, <}$, $\mathcal{G}_{\geq, \geq}$ in the following sections.

3. THE CASE : $\ell(I_1) < \ell(I_2)$ AND $\ell(J_1) < \ell(J_2)$

For the sake of convenience, we first present two key lemmas, which will be used later.

Lemma 3.1 ([8], [12]). *Let*

$$A_{I_1 I_2} = \frac{\ell(I_1)^{\alpha/2} \ell(I_2)^{\alpha/2}}{D(I_1, I_2)^{n+\alpha}} |I_1|^{1/2} |I_2|^{1/2},$$

where the long distance $D(I_1, I_2) = \ell(I_1) + \ell(I_2) + d(I_1, I_2)$, $I_1, I_2 \in \mathcal{D}_n$ and $\alpha > 0$. Then, for any $x_{I_1}, y_{I_2} \geq 0$, we have the following estimate,

$$\left(\sum_{I_1, I_2} A_{I_1 I_2} x_{I_1} y_{I_2} \right)^2 \lesssim \sum_{I_1} x_{I_1}^2 \times \sum_{I_2} y_{I_2}^2.$$

In particular, there holds that:

$$\sum_{I_2} \left[\sum_{I_1} A_{I_1 I_2} x_{I_1} \right]^2 \lesssim \sum_{I_1} x_{I_1}^2.$$

□

Lemma 3.2. *Let $0 < \alpha \leq n(\lambda_1 - 2)/2$. For a given cube $I_2 \in \mathcal{D}_n$ and $(x_1, t_1) \in W_{I_2}$, the following inequality holds,*

$$\left[\int_{\mathbb{R}^n} \left(\int_{I_1} \frac{dz_1}{(t_1 + |x_1 - y_1 - z_1|)^{n+\alpha}} \right)^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \frac{dy_1}{t_1^n} \right]^{1/2} \lesssim \frac{|I_1|}{(\ell(I_2) + d(I_1, I_2))^{n+\alpha}}.$$

Proof. Fixed $(x_1, t_1) \in W_{I_1}$. If $|y_1| \leq \frac{1}{2}d(I_1, I_2)$, then

$$t_1 + |x_1 - y_1 - z_1| \gtrsim \ell(I_2) + |x_1 - z_1| - |y_1| \gtrsim \ell(I_2) + d(I_1, I_2).$$

Thus, it follows that

$$\left[\int_{|y_1| \leq \frac{1}{2}d(I_1, I_2)} \left(\int_{I_1} \frac{dz_1}{(t_1 + |x_1 - y_1 - z_1|)^{n+\alpha}} \right)^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \frac{dy_1}{t_1^n} \right]^{1/2} \lesssim \frac{|I_1|}{(\ell(I_2) + d(I_1, I_2))^{n+\alpha}}.$$

If $|y_1| > \frac{1}{2}d(I_1, I_2)$, then

$$\left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \frac{1}{t_1^n} \lesssim \frac{\ell(I_2)^{n\lambda_1 - n}}{(\ell(I_2) + d(I_1, I_2))^{n\lambda_1}}.$$

Hence, we have

$$\begin{aligned} & \left[\int_{|y_1| > \frac{1}{2}d(I_1, I_2)} \left(\int_{I_1} \frac{dz_1}{(t_1 + |x_1 - y_1 - z_1|)^{n+\alpha}} \right)^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \frac{dy_1}{t_1^n} \right]^{1/2} \\ & \lesssim \frac{\ell(I_2)^{\frac{n\lambda_1}{2} - \frac{n}{2}}}{(\ell(I_2) + d(I_1, I_2))^{\frac{n\lambda_1}{2}}} \|\psi_{t_1} * \mathbf{1}_{I_1}\|_{L^2(\mathbb{R}^n)} \lesssim \frac{\ell(I_2)^{\frac{n\lambda_1}{2} - \frac{n}{2}}}{(\ell(I_2) + d(I_1, I_2))^{\frac{n\lambda_1}{2}}} \|\psi_{t_1}\|_{L^2(\mathbb{R}^n)} |I_1| \\ & \lesssim \frac{\ell(I_2)^{\frac{n\lambda_1}{2} - n - \alpha}}{(\ell(I_2) + d(I_1, I_2))^{\frac{n\lambda_1}{2}}} |I_1| \lesssim \frac{|I_1|}{(\ell(I_2) + d(I_1, I_2))^{n+\alpha}}. \end{aligned}$$

where $\psi_{t_1}(z_1) = (t_1 + |z_1|)^{-n-\alpha}$ and we have used the condition $0 < \alpha \leq n(\lambda_1 - 2)/2$ in the last step. □

Now we turn our attention to the estimate of $\mathcal{G}_{<,<}$. An easy consequence of the Hölder estimates of the kernel K_{t_1,t_2} is that:

$$\begin{aligned} |\theta_{t_1,t_2} h_{I_1 \times J_1}(x-y)| &\lesssim |I_1|^{-1/2} \int_{I_1} \frac{\ell(I_1)^\alpha}{(t_1 + |x_1 - y_1 - z_1|)^{n+\alpha}} dz_1 \\ &\quad \times |J_1|^{-1/2} \int_{J_1} \frac{\ell(J_1)^\beta}{(t_2 + |x_2 - y_2 - z_2|)^{m+\beta}} dz_2. \end{aligned}$$

Moreover, by Lemma 3.2, we can obtain that

$$\begin{aligned} &\mathcal{P}(x, t) \\ &:= \left(\iint_{\mathbb{R}^{n+m}} |\theta_{t_1,t_2} h_{I_1 \times J_1}(x-y)|^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \right)^{1/2} \\ (3.1) \quad &\lesssim |I_1|^{-1/2} \left[\int_{\mathbb{R}^n} \left(\int_{I_1} \frac{\ell(I_1)^\alpha dz_1}{(t_1 + |x_1 - y_1 - z_1|)^{n+\alpha}} \right)^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \frac{dy_1}{t_1^n} \right]^{1/2} \\ &\quad \times |J_1|^{-1/2} \left[\int_{\mathbb{R}^m} \left(\int_{J_1} \frac{\ell(J_1)^\beta dz_2}{(t_2 + |x_2 - y_2 - z_2|)^{m+\beta}} \right)^2 \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_2}{t_2^m} \right]^{1/2} \\ &\lesssim \frac{\ell(I_1)^\alpha}{(\ell(I_2) + d(I_1, I_2))^{n+\alpha}} |I_1|^{1/2} \frac{\ell(J_1)^\beta}{(\ell(J_2) + d(J_1, J_2))^{m+\beta}} |J_1|^{1/2}. \end{aligned}$$

Since $\ell(I_1) < \ell(I_2)$ and $\ell(J_1) < \ell(J_2)$, then we get

$$\mathcal{P}(x, t) \lesssim A_{I_1 I_2} |I_2|^{-1/2} \cdot A_{J_1 J_2} |J_2|^{-1/2}.$$

Therefore, by Minkowski's inequality and Lemma 3.1, it now follows that

$$\begin{aligned} \mathcal{G}_{<,<} &\lesssim \sum_{I_2, J_2: \text{good}} \iint_{W_{J_2}} \iint_{W_{I_2}} \left[\sum_{\substack{\ell(I_1) < \ell(I_2) \\ \ell(J_1) < \ell(J_2)}} |f_{I_1 J_1}| \left(\iint_{\mathbb{R}^{n+m}} |\theta_{t_1,t_2} h_{I_1 \times J_1}(x-y)|^2 \right. \right. \\ &\quad \times \left. \left. \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \right)^{1/2} \right]^2 \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2} \\ &\lesssim \sum_{J_2} \sum_{I_2} \left[\sum_{I_1} A_{I_1 I_2} \sum_{J_1} A_{J_1 J_2} |f_{I_1 J_1}| \right]^2 \\ &\lesssim \sum_{J_2} \sum_{I_1} \left[\sum_{J_1} A_{J_1 J_2} |f_{I_1 J_1}| \right]^2 \\ &\lesssim \sum_{I_1} \sum_{J_1} |f_{I_1 J_1}|^2 = \|f\|_{L^2(\mathbb{R}^{n+m})}^2. \end{aligned}$$

4. THE CASE : $\ell(I_1) \geq \ell(I_2)$ AND $\ell(J_1) < \ell(J_2)$

In any case, we perform the splitting

$$\sum_{\ell(I_1) \geq \ell(I_2)} = \sum_{\substack{\ell(I_1) \geq \ell(I_2) \\ d(I_1, I_2) > \ell(I_2)^\gamma \ell(I_1)^{1-\gamma}}} + \sum_{\substack{\ell(I_1) > 2^r \ell(I_2) \\ d(I_1, I_2) \leq \ell(I_2)^\gamma \ell(I_1)^{1-\gamma}}} + \sum_{\substack{\ell(I_2) \leq \ell(I_1) \leq 2^r \ell(I_2) \\ d(I_1, I_2) \leq \ell(I_2)^\gamma \ell(I_1)^{1-\gamma}}}.$$

These three parts are called separated, nested and adjacent respectively. The term nested makes sense, since the summing conditions that I_2 is good actually imply that I_1 is the ancestor of I_2 . Thus, there holds

$$\mathcal{G}_{\geq, <} \lesssim \mathcal{G}_{sep, <} + \mathcal{G}_{nes, <} + \mathcal{G}_{adj, <},$$

where

$$\begin{aligned} \mathcal{G}_{sep, <} &= \sum_{I_2, J_2: \text{good}} \iint_{W_{J_2}} \iint_{W_{I_2}} \iint_{\mathbb{R}^{n+m}} \left| \sum_{\substack{I_1: \ell(I_1) \geq \ell(I_2) \\ d(I_1, I_2) > \ell(I_2)^{\gamma_n} \ell(I_1)^{1-\gamma_n}}} \sum_{J_1: \ell(J_1) < \ell(J_2)} f_{I_1 J_1} \right. \\ &\quad \times \left. \theta_{t_1, t_2} h_{I_1 \times J_1}(x - y) \right|^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2}, \\ \mathcal{G}_{nes, <} &= \sum_{I_2, J_2: \text{good}} \iint_{W_{J_2}} \iint_{W_{I_2}} \iint_{\mathbb{R}^{n+m}} \left| \sum_{\substack{I_1: \ell(I_1) > 2^r \ell(I_2) \\ d(I_1, I_2) \leq \ell(I_2)^{\gamma_n} \ell(I_1)^{1-\gamma_n}}} \sum_{J_1: \ell(J_1) < \ell(J_2)} f_{I_1 J_1} \right. \\ &\quad \times \left. \theta_{t_1, t_2} h_{I_1 \times J_1}(x - y) \right|^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2}, \end{aligned}$$

and

$$\begin{aligned} \mathcal{G}_{adj, <} &= \sum_{I_2, J_2: \text{good}} \iint_{W_{J_2}} \iint_{W_{I_2}} \iint_{\mathbb{R}^{n+m}} \left| \sum_{\substack{I_1: \ell(I_2) \leq \ell(I_1) \leq 2^r \ell(I_2) \\ d(I_1, I_2) \leq \ell(I_2)^{\gamma_n} \ell(I_1)^{1-\gamma_n}}} \sum_{J_1: \ell(J_1) < \ell(J_2)} f_{I_1 J_1} \right. \\ &\quad \times \left. \theta_{t_1, t_2} h_{I_1 \times J_1}(x - y) \right|^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2}. \end{aligned}$$

Now, we are in the position to estimate the above three terms, respectively.

4.1. Separated Part $\mathcal{G}_{sep, <}$. In this case, we note that the following inequality holds,

$$(4.1) \quad \frac{t_1^\alpha}{(\ell(I_2) + d(I_1, I_2))^{n+\alpha}} |I_1|^{1/2} \lesssim A_{I_1 I_2} |I_2|^{-1/2}.$$

Indeed, if $d(I_1, I_2) \geq \ell(I_1)$, then $D(I_1, I_2) \sim d(I_1, I_2)$. Therefore, we get

$$\frac{t_1^\alpha}{(\ell(I_2) + d(I_1, I_2))^{n+\alpha}} |I_1|^{1/2} \lesssim A_{I_1 I_2} |I_2|^{-1/2}.$$

If $d(I_1, I_2) < \ell(I_1)$, then $D(I_1, I_2) \sim \ell(I_1)$. Moreover, notice that $\gamma_n(n + \alpha) = \alpha/2$ and $d(I_1, I_2) > \ell(I_2)^{\gamma_n} \ell(I_1)^{1-\gamma_n}$, one may conclude that

$$\frac{t_1^\alpha}{(\ell(I_2) + d(I_1, I_2))^{n+\alpha}} |I_1|^{1/2} \lesssim \frac{\ell(I_1)^{\alpha/2} \ell(I_2)^{\alpha/2}}{\ell(I_1)^{n+\alpha}} |I_1|^{1/2} \lesssim A_{I_1 I_2} |I_2|^{-1/2}.$$

It is obvious that the mixed Hölder and size condition implies that

$$\begin{aligned} |\theta_{t_1, t_2} h_{I_1 \times J_1}(x - y)| &\lesssim |I_1|^{-1/2} \int_{I_1} \frac{t_1^\alpha}{(t_1 + |x_1 - y_1 - z_1|)^{n+\alpha}} dz_1 \\ &\quad \times |J_1|^{-1/2} \int_{J_1} \frac{\ell(J_1)^\beta}{(t_2 + |x_2 - y_2 - z_2|)^{m+\beta}} dz_2. \end{aligned}$$

Thus, combining Lemma 3.2 with (4.1), one can obtain

$$\begin{aligned}
(4.2) \quad \mathcal{P}(x, t) &\lesssim |I_1|^{-1/2} \left[\int_{\mathbb{R}^n} \left(\int_{I_1} \frac{t_1^\alpha dz_1}{(t_1 + |x_1 - y_1 - z_1|)^{n+\alpha}} \right)^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \frac{dy_1}{t_1^n} \right]^{1/2} \\
&\quad \times |J_1|^{-1/2} \left[\int_{\mathbb{R}^m} \left(\int_{J_1} \frac{\ell(J_1)^\beta dz_2}{(t_2 + |x_2 - y_2 - z_2|)^{m+\beta}} \right)^2 \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_2}{t_2^m} \right]^{1/2} \\
&\lesssim \frac{t_1^\alpha}{(\ell(I_2) + d(I_1, I_2))^{n+\alpha}} |I_1|^{1/2} \frac{\ell(J_1)^\beta}{(\ell(J_2) + d(J_1, J_2))^{m+\beta}} |J_1|^{1/2} \\
&\lesssim \frac{\ell(I_2)^\alpha}{d(I_1, I_2)^{n+\alpha}} |I_1|^{1/2} A_{J_1 J_2} |J_2|^{-1/2} \lesssim A_{I_1 I_2} |I_2|^{-1/2} \cdot A_{J_1 J_2} |J_2|^{-1/2}.
\end{aligned}$$

Consequently, by the similar argument as $\mathcal{G}_{<, <}$, we have

$$\mathcal{G}_{sep, <} \lesssim \|f\|_{L^2(\mathbb{R}^{n+m})}^2.$$

4.2. Adjacent Part $\mathcal{G}_{adj, <}$. The summation conditions $\ell(I_2) \leq \ell(I_1) \leq 2^r \ell(I_2)$ and $d(I_1, I_2) \leq \ell(I_2)^{\gamma_n} \ell(I_1)^{1-\gamma_n}$ indicate that $\ell(I_1) \sim \ell(I_2) \sim D(I_1, I_2)$. Thus,

$$\frac{\ell(I_2)^\alpha}{(\ell(I_2) + d(I_1, I_2))^{n+\alpha}} |I_1|^{1/2} \lesssim \ell(I_2)^{-n} \sim \frac{\ell(I_1)^{\alpha/2} \ell(I_2)^{\alpha/2}}{D(I_1, I_2)^{n+\alpha}}.$$

It follows from (4.2) that

$$\mathcal{P}(x, t) \lesssim \frac{\ell(I_2)^\alpha}{(\ell(I_2) + d(I_1, I_2))^{n+\alpha}} |I_1|^{1/2} \cdot A_{J_1 J_2} |J_2|^{-1/2} \lesssim A_{I_1 I_2} |I_2|^{-1/2} \cdot A_{J_1 J_2} |J_2|^{-1/2}.$$

Therefore, exactly as we have seen before,

$$\mathcal{G}_{adj, <} \lesssim \|f\|_{L^2(\mathbb{R}^{n+m})}^2.$$

4.3. Nested Part $\mathcal{G}_{nes, <}$. We use $I^{(k)} \in \mathcal{D}_n$ to denote the unique cube for which $\ell(I^{(k)}) = 2^k \ell(I)$ and $I \subset I^{(k)}$. We call $I^{(k)}$ as the k generation older dyadic ancestor of I . In this case, by the goodness of I_2 , it must actually have $I_2 \subsetneq I_1$. That is, I_1 is the ancestor of I_2 . This enables us to write

$$\begin{aligned}
\mathcal{G}_{nes, <} &= \sum_{I, J_2: \text{good}} \iint_{W_{J_2}} \iint_{W_I} \iint_{\mathbb{R}^{n+m}} \left| \sum_{k=1}^{\infty} \sum_{J_1: \ell(J_1) < \ell(J_2)} f_{I^{(k)} J_1} \theta_{t_1, t_2} h_{I^{(k)} \times J_1}(x - y) \right|^2 \\
&\quad \times \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2}.
\end{aligned}$$

We introduce the notation

$$s_I^k = -\mathbf{1}_{(I^{(k-1)})^c} \langle h_{I^{(k)}} \rangle_{I^{(k-1)}} + \sum_{\substack{I' \in ch(I^{(k)}) \\ I' \neq I^{(k-1)}}} \mathbf{1}_{I'} h_{I^{(k)}}.$$

Then, it is easy to check that

$$(4.3) \quad h_{I^{(k)}} = s_I^k + \langle h_{I^{(k)}} \rangle_{I^{(k-1)}},$$

$\text{supp } s_I^k \subset (I^{(k-1)})^c$, and $|s_I^k| \lesssim |I^{(k)}|^{-1/2}$.

Denote $f_{J_1} = \langle f, h_{J_1} \rangle$ so that $f_{J_1} = \int_{\mathbb{R}^m} f(y_1, y_2) dy_2$, $y_1 \in \mathbb{R}^n$. Then, we split

$$\mathcal{G}_{nes, <} \lesssim \mathcal{G}_{mod, <} + \mathcal{G}_{Car, <} \quad ,$$

where

$$\begin{aligned} \mathcal{G}_{mod, <} &= \sum_{I, J_2: \text{good}} \iint_{W_{J_2}} \iint_{W_I} \iint_{\mathbb{R}^{n+m}} \left| \sum_{k=1}^{\infty} \sum_{J_1: \ell(J_1) < \ell(J_2)} f_{I^{(k)} J_1} \theta_{t_1, t_2}(s_I^k \otimes h_{J_1})(x - y) \right|^2 \\ &\quad \times \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2} \end{aligned}$$

and

$$\begin{aligned} \mathcal{G}_{Car, <} &= \sum_{I, J_2: \text{good}} \iint_{W_{J_2}} \iint_{W_I} \iint_{\mathbb{R}^{n+m}} \left| \sum_{J_1: \ell(J_1) < \ell(J_2)} f_{I^{(k)} J_1} \theta_{t_1, t_2}(\mathbf{1} \otimes h_{J_1})(x - y) \right. \\ &\quad \times \sum_{k=1}^{\infty} \langle \Delta_{I^{(k)}} f_{J_1} \rangle_{I^{(k-1)}} \left. \right|^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2} \\ &= \sum_{I, J_2: \text{good}} \iint_{W_{J_2}} \iint_{W_I} \iint_{\mathbb{R}^{n+m}} \left| \sum_{J_1: \ell(J_1) < \ell(J_2)} \langle f_{J_1} \rangle_I \theta_{t_1, t_2}(\mathbf{1} \otimes h_{J_1})(x - y) \right|^2 \\ &\quad \times \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2}. \end{aligned}$$

• **Estimate of $\mathcal{G}_{mod, <}$.** We need the following lemma.

Lemma 4.1. *Let $0 < \beta \leq m(\lambda_2 - 2)/2$ and $k \in \mathbb{N}_+$. Given cubes $I \in \mathcal{D}_n$, $J_1, J_2 \in \mathcal{D}_m$, $(x_1, t_1) \in W_I$, and $(x_2, t_2) \in W_{J_1}$, the following estimate holds*

$$\begin{aligned} \mathcal{Q}(x, t) &:= \left(\iint_{\mathbb{R}^{n+m}} |\theta_{t_1, t_2}(s_I^k \otimes h_{J_1})(x - y)|^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \right)^{1/2} \\ &\lesssim 2^{-\alpha k/2} |I^{(k)}|^{-1/2} |J_1|^{1/2} \frac{\ell(J_1)^\beta}{(\ell(J_2) + d(J_1, J_2))^{m+\beta}} \end{aligned}$$

Proof. By using the mixed Hölder and size condition, it yields that

$$\begin{aligned} |\theta_{t_1, t_2}(s_I^k \otimes h_{J_1})(x - y)| &\lesssim |I^{(k)}|^{-1/2} \int_{(I^{(k-1)})^c} \frac{t_1^\alpha}{(t_1 + |x_1 - y_1 - z_1|)^{n+\alpha}} dz_1 \\ &\quad \times |J_1|^{-1/2} \int_{J_1} \frac{\ell(J_1)^\beta}{(t_2 + |x_2 - y_2 - z_2|)^{m+\beta}} dz_2 \end{aligned}$$

Similarly as in (3.1), we only need to show

$$(4.4) \quad \mathcal{K} := \left[\int_{\mathbb{R}^n} \left(\int_{(I^{(k-1)})^c} \frac{t_1^\alpha dz_1}{(t_1 + |x_1 - y_1 - z_1|)^{n+\alpha}} \right)^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \frac{dy_1}{t_1^n} \right]^{1/2} \sim 2^{-\alpha k/2}.$$

Indeed, if $k \leq r$,

$$\begin{aligned} \mathcal{K} &\lesssim \left[\int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} \frac{\ell(I)^\alpha dz_1}{(\ell(I) + |x_1 - y_1 - z_1|)^{n+\alpha}} \right)^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \frac{dy_1}{t_1^n} \right]^{1/2} \\ &\lesssim \int_{\mathbb{R}^n} \frac{\ell(I)^\alpha}{(\ell(I) + |x_1 - z_1|)^{n+\alpha}} dz_1 \\ &\lesssim \ell(I)^{-n} |I| + \ell(I)^\alpha \int_{I^c} \frac{dz_1}{|z_1 - x_1|^{n+\alpha}} \lesssim 1 \sim 2^{-\alpha k/2}. \end{aligned}$$

If $k > r$, the goodness of I gives that

$$d(I, (I^{(k-1)})^c) > \ell(I)^{\gamma_n} \ell(I^{(k-1)})^{1-\gamma_n} = 2^{(k-1)(1-\gamma_n)} \ell(I) \gtrsim 2^{k/2} \ell(I).$$

Therefore, we obtain

$$\begin{aligned} \int_{(I^{(k-1)})^c} \frac{\ell(I)^\alpha}{|z_1 - x_1|^{n+\alpha}} dz_1 &\leq \int_{B(x_1, d(I, (I^{(k-1)})^c))} \frac{\ell(I)^\alpha}{|z_1 - x_1|^{n+\alpha}} dz_1 \\ &\lesssim \ell(I)^\alpha d(I, (I^{(k-1)})^c)^{-\alpha} \lesssim 2^{-\alpha k/2}. \end{aligned}$$

Given $y_1 \in \mathbb{R}^n$, we introduce the notation

$$E_1 = \{z_1 \in (I^{(k-1)})^c; |z_1 - x_1| \geq 2|y_1|\}, \quad E_2 = \{z_1 \in (I^{(k-1)})^c; |z_1 - x_1| < 2|y_1|\}.$$

Then, it follows that

$$\begin{aligned} \mathcal{K} &\lesssim \left[\int_{\mathbb{R}^n} \left(\int_{E_1} \frac{\ell(I)^\alpha dz_1}{(\ell(I) + |x_1 - y_1 - z_1|)^{n+\alpha}} \right)^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \frac{dy_1}{t_1^n} \right]^{1/2} \\ &\quad + \left[\int_{\mathbb{R}^n} \left(\int_{E_2} \frac{\ell(I)^\alpha dz_1}{(\ell(I) + |x_1 - y_1 - z_1|)^{n+\alpha}} \right)^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \frac{dy_1}{t_1^n} \right]^{1/2} \\ &:= \mathcal{K}_1 + \mathcal{K}_2. \end{aligned}$$

Note that $\ell(I) + |x_1 - y_1 - z_1| > |z_1 - x_1| - |y_1| \geq \frac{1}{2}|z_1 - x_1|$ whenever $z_1 \in E_1$. This yields that

$$\mathcal{K}_1 \lesssim \int_{(I^{(k-1)})^c} \frac{\ell(I)^\alpha}{|z_1 - x_1|^{n+\alpha}} dz_1 \cdot \left[\int_{\mathbb{R}^n} \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \frac{dy_1}{t_1^n} \right]^{1/2} \lesssim 2^{-\alpha k/2}.$$

As for \mathcal{K}_2 , let $\xi(z_1) = \frac{1}{t_1 + |z_1|^{n+\alpha}}$, $\eta(z_1) = \frac{\mathbf{1}_{E_2}(z_1)}{t_1 + |z_1 - x_1|^{n+\alpha}}$. By Young's inequality, we have

$$\begin{aligned} \mathcal{K}_2 &\lesssim \ell(I)^\alpha \left[\int_{\mathbb{R}^n} \left(\int_{E_2} \frac{\ell(I)^{\frac{n\lambda_1}{2} - \frac{n}{2}}}{(\ell(I) + |x_1 - z_1|)^{\frac{n\lambda_1}{2}}} \frac{1}{(t_1 + |y_1 - z_1|)^{n+\alpha}} dz_1 \right)^2 dy_1 \right]^{1/2} \\ &\leq \ell(I)^\alpha \left[\int_{\mathbb{R}^n} \left(\int_{E_2} \frac{\ell(I)^{\frac{n}{2} + \alpha}}{(\ell(I) + |x_1 - z_1|)^{n+\alpha}} \frac{1}{(t_1 + |y_1 - z_1|)^{n+\alpha}} dz_1 \right)^2 dy_1 \right]^{1/2} \\ &= \ell(I)^{\frac{n}{2} + 2\alpha} \|\xi * \eta\|_{L^2(\mathbb{R}^n)} \leq \ell(I)^{\frac{n}{2} + 2\alpha} \|\xi\|_{L^2(\mathbb{R}^n)} \|\eta\|_{L^1(\mathbb{R}^n)} \\ &\lesssim \ell(I)^\alpha \int_{(I^{(k-1)})^c} \frac{\ell(I)^\alpha}{|z_1 - x_1|^{n+\alpha}} dz_1 \lesssim 2^{-\alpha k/2}, \end{aligned}$$

□

Thus, by Minkowski's inequality and Lemma 4.1, $\mathcal{G}_{mod,<}$ can be controlled by

$$\begin{aligned}
& \mathcal{G}_{mod,<} \\
& \lesssim \sum_{I,J_2: \text{good}} \iint_{W_{J_2}} \iint_{W_I} \left[\sum_{k=1}^{\infty} \sum_{J_1: \ell(J_1) < \ell(J_2)} |f_{I^{(k)}J_1}| \left(\iint_{\mathbb{R}^{n+m}} |\theta_{t_1,t_2}(s_I^k \otimes h_{J_1})(x-y)|^2 \right. \right. \\
& \quad \left. \left. \times \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \right)^{1/2} \right]^2 \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2} \\
& \leq \sum_I \sum_{J_2} \iint_{W_{J_2}} \iint_{W_I} \left[\sum_{k=1}^{\infty} \frac{1}{2^{\alpha k/2}} \sum_{J_1: \ell(J_1) < \ell(J_2)} A_{J_1 J_2} |f_{I^{(k)}J_1}| \left(\frac{|J_1|}{|I^{(k)}|} \right)^{1/2} \right]^2 \frac{dx_1 dt_1 dx_2 dt_2}{t_1 t_2} \\
& \lesssim \sum_I \sum_{J_2} \left[\sum_{k=1}^{\infty} 2^{-\alpha k/2} \left(\frac{|I|}{|I^{(k)}|} \sum_{J_1: \ell(J_1) < \ell(J_2)} A_{J_1 J_2} |f_{I^{(k)}J_1}| \right)^{1/2} \right]^2 \\
& \leq \left[\sum_{k=1}^{\infty} 2^{-\alpha k/4} \cdot 2^{-\alpha k/4} \left(\sum_I \frac{|I|}{|I^{(k)}|} \sum_{J_2} \left(\sum_{J_1: \ell(J_1) < \ell(J_2)} A_{J_1 J_2} |f_{I^{(k)}J_1}| \right)^2 \right)^{1/2} \right]^2 \\
& \lesssim \sum_{k=1}^{\infty} 2^{-\alpha k/2} \sum_I \frac{|I|}{|I^{(k)}|} \sum_{J_2} \left(\sum_{J_1: \ell(J_1) < \ell(J_2)} A_{J_1 J_2} |f_{I^{(k)}J_1}| \right)^2 \\
& \lesssim \sum_{k=1}^{\infty} 2^{-\alpha k/2} \sum_{Q, J_1} \frac{|f_{QJ_1}|^2}{|Q|} \sum_{I: I^{(k)}=Q} |I| \lesssim \|f\|_{L^2(\mathbb{R}^{n+m})}^2.
\end{aligned}$$

□

• **Estimate of $\mathcal{G}_{Car,<}$.** We need to use the following lemma.

Lemma 4.2. *Let $J_1, J_2 \in \mathcal{D}_m$ be cubes, and $(x_2, t_2) \in W_{J_2}$. Then the Carleson condition holds*

$$\begin{aligned}
\mathcal{R}(x_2, t_2) &:= \sum_{I' \subset I} \iint_{W_{I'}} \iint_{\mathbb{R}^{n+m}} |\theta_{t_1,t_2}(\mathbf{1} \otimes h_{J_1})(x-y)|^2 \prod_{i=1}^2 \left(\frac{t_i}{t_i + |y_i|} \right)^{n\lambda_i} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \\
&\lesssim |I| \left(\frac{\ell(J_1)^\beta |J_1|^{1/2}}{(\ell(J_2) + d(J_1, J_2))^{m+\beta}} \right)^2.
\end{aligned}$$

Proof. The first step is to split

$$\mathcal{R}(x_2, t_2) \lesssim \mathcal{R}_1(x_2, t_2) + \mathcal{R}_2(x_2, t_2),$$

where

$$\mathcal{R}_1(x_2, t_2) = \iint_{\widehat{3I}} \iint_{\mathbb{R}^{n+m}} |\theta_{t_1,t_2}(\mathbf{1}_{3I} \otimes h_{J_1})(x-y)|^2 \prod_{i=1}^2 \left(\frac{t_i}{t_i + |y_i|} \right)^{n\lambda_i} \frac{dy_1 dy_2}{t_1^n t_2^m} \frac{dx_1 dt_1}{t_1},$$

and

$$\begin{aligned}\mathcal{R}_2(x_2, t_2) &= \iint_{\widehat{I}} \iint_{\mathbb{R}^{n+m}} |\theta_{t_1, t_2}(\mathbf{1}_{(3I)^c} \otimes h_{J_1})(x - y)|^2 \prod_{i=1}^2 \left(\frac{t_i}{t_i + |y_i|} \right)^{n\lambda_i} \frac{dy_1 dy_2}{t_1^n t_2^m} \frac{dx_1 dt_1}{t_1} \\ &:= \iint_{\widehat{I}} H(x, t) \frac{dx_1 dt_1}{t_1}.\end{aligned}$$

By the combinations of Carleson and Hölder conditions and Lemma 3.2, it follows that

$$\begin{aligned}\mathcal{R}_1(x_2, t_2) &\lesssim \frac{1}{|J_1|} \int_{\mathbb{R}^m} \iint_{\widehat{3I}} \int_{\mathbb{R}^n} \left| \int_{J_1} \int_{3I} [K_{t_1, t_2}(x - y, (z_1, z_2)) - K_{t_1, t_2}(x - y, (z_1, z_2 + c_{J_1}))] dz_1 dz_2 \right|^2 \\ &\quad \times \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \frac{dy_1}{t_1^n} \frac{dx_1 dt_1}{t_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_2}{t_2^m} \\ &\quad \frac{1}{|J_1|} \int_{\mathbb{R}^m} \left[\int_{J_1} \left(\iint_{\widehat{3I}} \int_{\mathbb{R}^n} \left| \int_{3I} [K_{t_1, t_2}(x - y, (z_1, z_2)) - K_{t_1, t_2}(x - y, (z_1, z_2 + c_{J_1}))] dz_1 \right|^2 \right. \right. \\ &\quad \left. \left. \times \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \frac{dy_1}{t_1^n} \frac{dx_1 dt_1}{t_1} \right)^{1/2} dz_2 \right]^2 \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_2}{t_2^m} \\ &\lesssim |I| |J_1|^{-1} \int_{\mathbb{R}^m} \left(\int_{J_1} \frac{\ell(J_1)^\beta dz_2}{(t_2 + |x_2 - y_2 - z_2|)^{m+\beta}} \right)^2 \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_2}{t_2^m} \\ &\lesssim |I| \left(\frac{\ell(J_1)^\beta |J_1|^{1/2}}{(\ell(J_2) + d(J_1, J_2))^{m+\beta}} \right)^2.\end{aligned}$$

The mixed Hölder and size estimate gives that

$$\begin{aligned}|\theta_{t_1, t_2}(\mathbf{1}_{(3I)^c} \otimes h_{J_1})(x - y)| &\lesssim |J_1|^{-1/2} \int_{(3I)^c} \frac{t_1^\alpha}{(t_1 + |x_1 - y_1 - z_1|)^{n+\alpha}} dz_1 \\ &\quad \times \int_{J_1} \frac{\ell(J_1)^\beta}{(t_2 + |x_2 - y_2 - z_2|)^{m+\beta}} dz_2.\end{aligned}$$

Thus, by the estimates in Lemma 3.2 and (4.4), one can deduce that

$$\begin{aligned}H(x, t) &\lesssim |J_1|^{-1} \int_{\mathbb{R}^n} \left(\int_{(3I)^c} \frac{t_1^\alpha dz_1}{(t_1 + |x_1 - y_1 - z_1|)^{n+\alpha}} \right)^2 \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \frac{dy_1}{t_1^n} \\ &\quad \times \int_{\mathbb{R}^m} \left(\int_{J_1} \frac{\ell(J_1)^\beta dz_2}{(t_2 + |x_2 - y_2 - z_2|)^{m+\beta}} \right)^2 \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_2}{t_2^m} \\ &\lesssim t_1^{2\alpha} \ell(I)^{-2\alpha} \left(\frac{\ell(J_1)^\beta |J_1|^{1/2}}{(\ell(J_2) + d(J_1, J_2))^{m+\beta}} \right)^2.\end{aligned}$$

Therefore, we obtain

$$\begin{aligned}\mathcal{R}_2(x_2, t_2) &= \iint_{\hat{I}} H(x, t) \frac{dx_1 dt_1}{t_1} \\ &\lesssim |I| \ell(I)^{-2\alpha} \int_0^{\ell(I)} t_1^{2\alpha-1} dt_1 \cdot \left(\frac{\ell(J_1)^\beta |J_1|^{1/2}}{(\ell(J_2) + d(J_1, J_2))^{m+\beta}} \right)^2 \\ &\lesssim |I| \left(\frac{\ell(J_1)^\beta |J_1|^{1/2}}{(\ell(J_2) + d(J_1, J_2))^{m+\beta}} \right)^2.\end{aligned}$$

Thus, we finish the proof of Lemma 4.2. □

Now we give the estimate for $\mathcal{G}_{Car, <}$. If $\ell(J_1) < \ell(J_2)$, then we have

$$\mathcal{R}(x_2, t_2) \lesssim |I| (A_{J_1, J_2} |J_2|^{-1/2})^2.$$

Therefore, we obtain the following estimate

$$\begin{aligned}\mathcal{G}_{Car, <} &= \sum_{J_2: \text{good}} \iint_{W_{J_2}} \sum_{I: \text{good}} \iint_{W_I} \iint_{\mathbb{R}^{n+m}} \left| \sum_{J_1: \ell(J_1) < \ell(J_2)} \langle f_{J_1} \rangle_I \theta_{t_1, t_2}(\mathbf{1} \otimes h_{J_1})(x - y) \right|^2 \\ &\quad \times \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2} \\ &\leq \sum_{J_2} \iint_{W_{J_2}} \sum_I \left[\sum_{J_1: \ell(J_1) < \ell(J_2)} \left(\iint_{W_I} \iint_{\mathbb{R}^{n+m}} \left| \langle f_{J_1} \rangle_I \theta_{t_1, t_2}(\mathbf{1} \otimes h_{J_1})(x - y) \right|^2 \right. \right. \\ &\quad \times \left. \left. \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \right)^{1/2} \right]^2 \frac{dx_2 dt_2}{t_2} \\ &\leq \sum_{J_2} \iint_{W_{J_2}} \left[\sum_{J_1: \ell(J_1) < \ell(J_2)} \left(\sum_I |\langle f_{J_1} \rangle_I|^2 \iint_{W_I} \iint_{\mathbb{R}^{n+m}} \left| \theta_{t_1, t_2}(\mathbf{1} \otimes h_{J_1})(x - y) \right|^2 \right. \right. \\ &\quad \times \left. \left. \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \right)^{1/2} \right]^2 \frac{dx_2 dt_2}{t_2} \\ &\lesssim \sum_{J_2} \left[\sum_{J_1: \ell(J_1) < \ell(J_2)} A_{J_1 J_2} \|f_{J_1}\|_{L^2(\mathbb{R}^n)} \right]^2 \lesssim \sum_{J_1} \|f_{J_1}\|_{L^2(\mathbb{R}^n)}^2 \lesssim \|f\|_{L^2(\mathbb{R}^{n+m})}^2.\end{aligned}$$

So far, we have completed the estimate of $\mathcal{G}_{\geq, <}$.

As for the term $\mathcal{G}_{<, \geq}$, it is completely symmetric with the term $\mathcal{G}_{\geq, <}$. It is worth noting that the mixed Hölder and size estimate and the combination of Carleson and Hölder estimate are symmetric, respectively. Thus the estimate for $\mathcal{G}_{<, \geq}$ is also true and we here omit its proof.

5. THE CASE : $\ell(I_1) \geq \ell(I_2)$ AND $\ell(J_1) \geq \ell(J_2)$.

Similar as what we have done before, the summation $\ell(I_1) \geq \ell(I_2)$ was decomposed into the separated, nested and adjacent terms. A similar splitting in the summation

$\ell(J_1) \geq \ell(J_2)$ is also performed. This splits the whole summation into nine parts as follows.

$$\begin{aligned} \mathcal{G}_{\geq, \geq} &\lesssim \mathcal{G}_{sep, sep} + \mathcal{G}_{sep, nes} + \mathcal{G}_{sep, adj} + \mathcal{G}_{nes, sep} + \mathcal{G}_{nes, nes} \\ &\quad + \mathcal{G}_{nes, adj} + \mathcal{G}_{adj, sep} + \mathcal{G}_{adj, nes} + \mathcal{G}_{adj, adj}. \end{aligned}$$

5.1. Nested/Nested : $\mathcal{G}_{nes, nes}$. We begin with the term $\mathcal{G}_{nes, nes}$, where the new bi-parameter phenomena will appear. Note that although this is only one of the many cases one needs to discuss in order to obtain a full estimate for $\mathcal{G}_{\geq, \geq}$ term, all the main difficulties in other cases are in fact already embedded in Nested/Nested. The fact will become more and more clear throughout the proof. Similarly, for the singular integral operators including bi-parameter and multi-parameter cases, the Nested part is also the most difficult one. Because it involves in some paraproduct estimates and all the BMO type estimates.

The decomposition of $h_{I^{(k)}}$ in (4.3) gives that

$$\mathcal{G}_{nes, nes} \lesssim \mathcal{G}_{mod, mod} + \mathcal{G}_{Car, Car} + \mathcal{G}_{mod, Car} + \mathcal{G}_{Car, mod},$$

where

$$\begin{aligned} \mathcal{G}_{mod, mod} &= \sum_{I, J: good} \iint_{W_J} \iint_{W_I} \iint_{\mathbb{R}^{n+m}} \left| \sum_{k=1}^{\infty} \sum_{i=1}^{\infty} f_{I^{(k)}} \theta_{t_1, t_2} (s_I^k \otimes s_J^i)(x-y) \right|^2 \\ &\quad \times \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2}, \end{aligned}$$

$$\begin{aligned} \mathcal{G}_{mod, Car} &= \sum_{I, J: good} \iint_{W_J} \iint_{W_I} \iint_{\mathbb{R}^{n+m}} \left| \sum_{k=1}^{\infty} \langle f_{I^{(k)}} \rangle_J \theta_{t_1, t_2} (s_I^k \otimes \mathbf{1})(x-y) \right|^2 \\ &\quad \times \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2}, \end{aligned}$$

$$\begin{aligned} \mathcal{G}_{Car, mod} &= \sum_{I, J: good} \iint_{W_J} \iint_{W_I} \iint_{\mathbb{R}^{n+m}} \left| \sum_{\ell=1}^{\infty} \langle f_{J^{(\ell)}} \rangle_I \theta_{t_1, t_2} (\mathbf{1} \otimes s_J^{\ell})(x-y) \right|^2 \\ &\quad \times \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2}, \end{aligned}$$

and

$$\begin{aligned} \mathcal{G}_{Car, Car} &= \sum_{I, J: good} |\langle f \rangle_{I \times J}|^2 \iint_{W_J} \iint_{W_I} \iint_{\mathbb{R}^{n+m}} |\theta_{t_1, t_2}(\mathbf{1})(x-y)|^2 \\ &\quad \times \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2}. \end{aligned}$$

5.1.1. **Estimate of $\mathcal{G}_{mod,mod}$.** We proceed using the standard argument as in Lemma 4.1. The size condition and (4.4) lead to the bound

$$\begin{aligned} \left(\iint_{\mathbb{R}^{n+m}} |\theta_{t_1,t_2}(s_I^k \otimes s_J^i)(x-y)|^2 \left(\frac{t_1}{t_1+|y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2+|y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \right)^{1/2} \\ \lesssim 2^{-\alpha k/2} |I^{(k)}|^{-1/2} \cdot 2^{-\beta i} |J^{(i)}|^{-1/2}. \end{aligned}$$

It is similar to estimate $\mathcal{G}_{mod,<}$ to analyze $\mathcal{G}_{mod,mod}$.

$$\begin{aligned} \mathcal{G}_{mod,mod} &\lesssim \sum_{k,i} 2^{-\alpha k/2} 2^{-\beta i} \sum_{Q,R} |f_{QR}|^2 \frac{1}{|Q|} \sum_{I:I^{(k)}=Q} |I| \cdot \frac{1}{|R|} \sum_{J:J^{(i)}=R} |J| \\ &\lesssim \|f\|_{L^2(\mathbb{R}^{n+m})}^2. \end{aligned}$$

5.1.2. **Estimate of $\mathcal{G}_{Car,Car}$.** Applying the bi-parameter Carleson condition, it immediately yields that

$$\begin{aligned} \mathcal{G}_{Car,Car} &= \sum_{I,J} |\langle f \rangle_{I \times J}|^2 C_{IJ}^{\mathcal{D}} = 2 \int_0^\infty \sum_{\substack{I,J \\ |\langle f \rangle_{I \times J}| > t}} C_{IJ}^{\mathcal{D}} t \, dt \\ &\lesssim \int_0^\infty \sum_{\substack{I,J \\ I \times J \subset \{M_{\mathcal{D}}f > t\}}} C_{IJ}^{\mathcal{D}} t \, dt \lesssim \int_0^\infty |\{M_{\mathcal{D}}f > t\}| t \, dt \\ &\lesssim \|M_{\mathcal{D}}f\|_{L^2(\mathbb{R}^{n+m})}^2 \lesssim \|f\|_{L^2(\mathbb{R}^{n+m})}^2, \end{aligned}$$

where in the last step we have used the $L^p(1 < p < \infty)$ boundedness of the strong maximal function associated with rectangles.

5.1.3. **Estimate of $\mathcal{G}_{Car,mod}$ and $\mathcal{G}_{mod,Car}$.**

Lemma 5.1. *Let $J \in \mathcal{D}_{m,good}$, $(x_2, t_2) \in W_{J_2}$ and $i \in \mathbb{N}$ be fixed. Then the Carleson condition is satisfied*

$$\begin{aligned} \sum_{I' \subset I} \iint_{W_{I'}} \iint_{\mathbb{R}^{n+m}} |\theta_{t_1,t_2}(\mathbf{1} \otimes s_J^i)(x-y)|^2 \left(\frac{t_1}{t_1+|y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2+|y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \\ \lesssim 2^{-\beta i} |I| \cdot |J^{(i)}|^{-1}. \end{aligned}$$

The proof of Lemma 5.1 is similar to Lemma 4.2. The size condition and mixed Carleson and size estimate are used. In addition, the inequality (4.4) is used twice. \square

Therefore, $\mathcal{G}_{Car,mod}$ is bounded as below.

$$\begin{aligned} \mathcal{G}_{Car,mod} &\leq \sum_{J:good} \iint_{W_J} \sum_I \left[\sum_{i=1}^{\infty} |\langle f_{J^{(i)}} \rangle_I| \left(\iint_{W_I} \iint_{\mathbb{R}^{n+m}} |\theta_{t_1,t_2}(\mathbf{1} \otimes s_J^i)(x-y)|^2 \right. \right. \\ &\quad \times \left. \left. \left(\frac{t_1}{t_1 + |y_1|} \right)^{n\lambda_1} \left(\frac{t_2}{t_2 + |y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \right)^{1/2} \right]^2 \frac{dx_2 dt_2}{t_2} \\ &\lesssim \sum_{i=1}^{\infty} 2^{-\beta i/2} \sum_R \|f_R\|_{L^2(\mathbb{R}^n)}^2 \frac{1}{|R|} \sum_{J:J^{(i)}=R} |I| \lesssim \|f\|_{L^2(\mathbb{R}^{n+m})}^2. \end{aligned}$$

5.2. The rest of terms. As for the estimates of the remaining terms, they are simply combinations of the techniques we have used above. Thereby, we here only present certain key points.

When reviewing the above proof, one will realize that the central part is to dominate $\mathcal{P}(x, t)$, $\mathcal{Q}(x, t)$ and $\mathcal{R}(x_2, t_2)$. So do the rest of terms. Moreover, the initial estimates of \mathcal{P} , \mathcal{Q} and \mathcal{R} are retained in the inequality (3.1), Lemma 4.1 and Lemma 4.2 respectively. They do not involve the relationship of side length of cubes I_1 , I_2 , J_1 and J_2 . Thus, based on the inequality (3.1), Lemma 4.1 and Lemma 4.2, one only needs to add the corresponding the relationship of side length.

Consequently, using the size condition or the mixed Hölder and size condition, it yields the bounds for $\mathcal{G}_{sep,sep}$, $\mathcal{G}_{sep,adj}$, $\mathcal{G}_{adj,adj}$ and $\mathcal{G}_{adj,sep}$ directly. Finally, for the terms $\mathcal{G}_{nes,sep}$ and $\mathcal{G}_{nes,adj}$, nes is split into mod and Car . Applying the size condition and the combinations of Carleson and size estimate, we will bound them. The terms $\mathcal{G}_{sep,nes}$ and $\mathcal{G}_{adj,nes}$ are symmetric with respect to them respectively. \square

6. THE NECESSITY OF BI-PARAMETER CARLESON CONDITION

We here show that the bi-parameter Carleson condition is necessary for $g_{\lambda_1, \lambda_2}^*$ -function to be bounded on $L^2(\mathbb{R}^{n+m})$.

Suppose that $\theta_{t_1, t_2} = \theta_{t_1}^n \otimes \theta_{t_2}^m$ is bounded on $L^2(\mathbb{R}^{n+m})$, where $\theta_{t_1}^n$ has a kernel $s_{t_1}^n(x_1, y_1)$, $\theta_{t_2}^m$ has a kernel $s_{t_2}^m(x_2, y_2)$, $x_1, y_1 \in \mathbb{R}^n$, $x_2, y_2 \in \mathbb{R}^m$, $t_1, t_2 > 0$. We assume that these satisfy the size condition and the corresponding L^2 bounds in \mathbb{R}^n and \mathbb{R}^m . We shall show that the bi-parameter Carleson condition (1.1) holds.

Define $\tilde{\Omega} = \{M_{\mathcal{D}} \mathbf{1}_{\Omega} > 1/2\}$ and $\hat{\Omega} = \{M \mathbf{1}_{\tilde{\Omega}} > c\}$ for a small enough dimensional constant $c = c(n, m)$, where $M_{\mathcal{D}}$ denote the strong maximal function related to the grid \mathcal{D} and M denote the strong maximal function. From the endpoint estimates for M and $M_{\mathcal{D}}$, it follows that $|\hat{\Omega}| \lesssim |\tilde{\Omega}| \lesssim |\Omega|$. Hence, it is enough to show that

$$\begin{aligned} &\sum_{\substack{I \times J \in \mathcal{D} \\ I \times J \subset \Omega}} \iint_{W_J} \iint_{W_I} \iint_{\mathbb{R}^{n+m}} |\theta_{t_1, t_2} \mathbf{1}_{\hat{\Omega}^c}(y_1, y_2)|^2 \left(\frac{t_1}{t_1 + |x_1 - y_1|} \right)^{n\lambda_1} \\ &\quad \times \left(\frac{t_2}{t_2 + |x_2 - y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2} \lesssim |\Omega|. \end{aligned}$$

For every $J \in \mathcal{D}_m$ we let \mathcal{F}_J consist of the maximal $F \in \mathcal{D}_n$ for which $F \times J \subset \tilde{\Omega}$. Then we define $F_J := \bigcup_{F \in \mathcal{F}_J} 2F$. Moreover, for fixed $I \in \mathcal{D}_n$, let \mathcal{G}_I be the family of the

maximal $G \in \mathcal{D}_m$ for which $I \times G \subset \Omega$, and $I_G \in \mathcal{D}_n$ be the maximal cube for which $I_G \supset I$ and $I_G \times G \subset \tilde{\Omega}$. So, we only need to show the following inequalities.

$$\begin{aligned} \mathcal{G}_1 &= \sum_{\substack{I \times J \in \mathcal{D} \\ I \times J \subset \Omega}} \iint_{W_J} \iint_{W_I} \iint_{\mathbb{R}^{n+m}} |\theta_{t_1, t_2}(\mathbf{1}_{\hat{\Omega}^c} \mathbf{1}_{F_J})(y_1, y_2)|^2 \left(\frac{t_1}{t_1 + |x_1 - y_1|} \right)^{n\lambda_1} \\ &\quad \times \left(\frac{t_2}{t_2 + |x_2 - y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2} \\ &:= \sum_J \iint_{W_J} \mathcal{G}_J(x_2, t_2) \frac{dx_2 dt_2}{t_2} \lesssim |\Omega|, \end{aligned}$$

and

$$\begin{aligned} \mathcal{G}_2 &= \sum_{\substack{I \times J \in \mathcal{D} \\ I \times J \subset \Omega}} \iint_{W_J} \iint_{W_I} \iint_{\mathbb{R}^{n+m}} |\theta_{t_1, t_2}(\mathbf{1}_{\hat{\Omega}^c} \mathbf{1}_{F_J^c})(y_1, y_2)|^2 \left(\frac{t_1}{t_1 + |x_1 - y_1|} \right)^{n\lambda_1} \\ &\quad \times \left(\frac{t_2}{t_2 + |x_2 - y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \frac{dx_2 dt_2}{t_2} \\ &:= \sum_I \iint_{W_I} \mathcal{G}_I(x_1, t_1) \frac{dx_1 dt_1}{t_1} \lesssim |\Omega|. \end{aligned}$$

To attain the goal, we need to first bound $\mathcal{G}_J(x_2, t_2)$ and $\mathcal{G}_I(x_1, t_1)$. Actually, Minkowski's integral inequality and size estimate yield that

$$\begin{aligned} \mathcal{G}_J(x_2, t_2) &\lesssim \left[\int_{\mathbb{R}^m} \left(\iint_{\mathbb{R}_+^{n+1}} \iint_{\mathbb{R}^{n+m}} |K_{t_2}^m(y_2, z_2)|^2 |\theta_{t_1, t_2}(\mathbf{1}_{\hat{\Omega}^c} \mathbf{1}_{F_J})(y_1)|^2 \left(\frac{t_1}{t_1 + |x_1 - y_1|} \right)^{n\lambda_1} \right. \right. \\ &\quad \left. \left. \times \left(\frac{t_2}{t_2 + |x_2 - y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_1 dt_1}{t_1} \right)^{1/2} dz_2 \right]^2 \\ &\lesssim \left[\int_{\mathbb{R}^m} \left(\int_{\mathbb{R}^m} \left(\frac{t_2^\beta}{(t_2 + |y_2 - z_2|)^{m+\beta}} \right)^2 \left(\frac{t_2}{t_2 + |x_2 - y_2|} \right)^{m\lambda_2} \frac{dy_2}{t_2^m} \right)^{1/2} \right. \\ &\quad \left. \times \left\| \mathbf{1}_{\hat{\Omega}^c}(\cdot, z_2) \mathbf{1}_{F_J} \right\|_{L^2(\mathbb{R}^n)} dz_2 \right]^2 \\ &\lesssim \left[\int_{\mathbb{R}^m} \frac{\ell(J)^\beta}{(\ell(J) + |x_2 - z_2|)^{m+\beta}} \left\| \mathbf{1}_{\hat{\Omega}^c}(\cdot, z_2) \mathbf{1}_{F_J} \right\|_{L^2(\mathbb{R}^n)} dz_2 \right]^2 \\ &\lesssim \int_{\mathbb{R}^m} \frac{\ell(J)^\beta}{(\ell(J) + |x_2 - z_2|)^{m+\beta}} \left\| \mathbf{1}_{\hat{\Omega}^c}(\cdot, z_2) \mathbf{1}_{F_J} \right\|_{L^2(\mathbb{R}^n)}^2 dz_2 \\ &\lesssim \int_{\mathbb{R}^m} \frac{\ell(J)^\beta}{|z_2 - c_J|^{m+\beta}} \left\| \mathbf{1}_{\hat{\Omega}^c}(\cdot, z_2) \mathbf{1}_{F_J} \right\|_{L^2(\mathbb{R}^n)}^2 dz_2 \\ &= \int_{\mathbb{R}^n} \mathbf{1}_{F_J}(z_1) \int_{\mathbb{R}^m} \frac{\ell(J)^\beta}{|z_2 - c_J|^{m+\beta}} \mathbf{1}_{\hat{\Omega}^c}(z_1, z_2) dz_2 dz_1. \end{aligned}$$

Similarly, we may estimate

$$\begin{aligned}
\mathcal{G}_I(x_1, t_1) &= \sum_{G \in \mathcal{G}_I} \sum_{J: J \subset G} \iint_{W_J} \iint_{\mathbb{R}^{n+m}} |\theta_{t_1, t_2}(\mathbf{1}_{\widehat{\Omega}^c} \mathbf{1}_{F_J^c})(y_1, y_2)|^2 \left(\frac{t_1}{t_1 + |x_1 - y_1|} \right)^{n\lambda_1} \\
&\quad \times \left(\frac{t_2}{t_2 + |x_2 - y_2|} \right)^{m\lambda_2} \frac{dy_1}{t_1^n} \frac{dy_2}{t_2^m} \frac{dx_2 dt_2}{t_2} \\
&\lesssim \left[\int_{\mathbb{R}^n} \frac{\ell(I)^\alpha}{(\ell(I) + |x_1 - z_1|)^{n+\alpha}} \left(\sum_{G \in \mathcal{G}_I} \mathbf{1}_{(2I_G)^c}(z_1) |G| \right)^{1/2} dz_1 \right]^2 \\
&\lesssim \sum_{G \in \mathcal{G}_I} |G| \int_{\mathbb{R}^n} \frac{\ell(I)^\alpha}{(\ell(I) + |x_1 - z_1|)^{n+\alpha}} \mathbf{1}_{(2I_G)^c}(z_1) dz_1 \\
&\lesssim \sum_{G \in \mathcal{G}_I} |G| \int_{I_G^c} \frac{\ell(I)^\alpha}{|z_1 - c_{I_G}|^{n+\alpha}} dz_1 \lesssim \sum_{G \in \mathcal{G}_I} |G| \left(\frac{\ell(I)}{\ell(I_G)} \right)^\alpha.
\end{aligned}$$

The remaining calculation is a routine application of the idea of [12]. We here omit the details. Finally, we obtain

$$\mathcal{G}_1 = \sum_J \iint_{W_J} \mathcal{G}_J(x_2, t_2) \frac{dx_2 dt_2}{t_2} \lesssim |\Omega|, \quad \mathcal{G}_2 = \sum_I \iint_{W_I} \mathcal{G}_I(x_1, t_1) \frac{dx_1 dt_1}{t_1} \lesssim |\Omega|.$$

Thus, we have proved the necessity. \square

REFERENCES

- [1] M. Cao, K. Li, Q. Xue, *A Characterization of two weight norm inequality for Littlewood-Paley g_λ^* -function*, available at <http://arxiv.org/abs/1504.07850>
- [2] M. Cao, Q. Xue, *A non-homogeneous local Tb theorem for Littlewood-Paley g_λ^* -function with L^p -testing condition*, available at <http://arxiv.org/abs/1507.05291>
- [3] C. Fefferman, *Inequalities for strongly singular convolution operators*, Acta. Math., 124 (1970), 9-36.
- [4] T. Hytönen, *The sharp weighted bound for general Calderon-Zygmund operators*, Ann. Math., (2) 175 (3) (2012), 1473-1506.
- [5] T. Hytönen, H. Martikainen, *Non-homogeneous T1 theorem for bi-parameter singular integrals*, Adv. Math., 261 (2014), 220-273.
- [6] M. T. Lacey, *The two weight inequality for the Hilbert transform: a primer*, submitted (2013), available at <http://www.arxiv.org/abs/1304.5004>.
- [7] M. T. Lacey, *Two weight inequality for the Hilbert transform: a real variable characterization, II*, Duke Math. J., 163 (2014), no. 15, 2821-2840.
- [8] M. T. Lacey, K. Li, *Two weight norm inequalities for g function*, Math. Res. Lett., 21 (2014), no. 03, 521-536.
- [9] J. Littlewood, R. Paley, *Theorems on Fourier series and power series, II*, Proc. Lond. Math. Soc., 42 (1936), 52-89.
- [10] M. T. Lacey, E. T. Sawyer, I. Uriarte-Tuero and C.-Y. Shen, *Two weight inequality for the Hilbert transform: a real variable characterization, I*, Duke Math. J., 163 (2014), no. 15, 2795-2820.
- [11] H. Martikainen, *Representation of bi-parameter singular integrals by dyadic operators*, Adv. Math., 229 (3) (2012), 1734-1761.
- [12] H. Martikainen, *Boundedness of a class of bi-parameter square function in the upper half-space*, J. Funct. Anal., 267 (2014), 3580-3597.

- [13] F. Nazarov, S. Treil and A. Volberg, *The Tb-theorem on non-homogeneous spaces*, Acta Math., 190 (2) (2003), 151-239.
- [14] Y. Ou, *A $T(b)$ theorem on product spaces*, Trans. Amer. Math. Soc., 367 (9)(2015), 6159-6197.
- [15] E. M. Stein, *On some function of Littlewood-Paley and Zygmund*, Bull. Amer. Math. Soc., 67 (1961), 99-101.
- [16] E. M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton Univ. Press, Princeton, NJ, 1970.

MINGMING CAO, SCHOOL OF MATHEMATICAL SCIENCES, BEIJING NORMAL UNIVERSITY, LABORATORY OF MATHEMATICS AND COMPLEX SYSTEMS, MINISTRY OF EDUCATION, BEIJING 100875, PEOPLE'S REPUBLIC OF CHINA

E-mail address: m.cao@mail.bnu.edu.cn

QINGYING XUE, SCHOOL OF MATHEMATICAL SCIENCES, BEIJING NORMAL UNIVERSITY, LABORATORY OF MATHEMATICS AND COMPLEX SYSTEMS, MINISTRY OF EDUCATION, BEIJING 100875, PEOPLE'S REPUBLIC OF CHINA

E-mail address: qyxue@bnu.edu.cn